DOCUMENT RESUME

ED 073 387 CG 007 819

AUTHOR Barnard, James D.; And Others
TITLE Supervising Paraprofessionals: Performance-Related Feedback.
INSTITUTION Kansas Univ., Lawrence. Dept. of Human Development.
SPONS AGENCY National Coordination Center for Early Childhood Education, St. Ann, Mo.; National Inst. of Child Health and Human Development (NIH), Bethesda, Md.
PUB DATE 72
GRANT OEG-3-7-070706-3118
NOTE 17p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Evaluation; *Feedback; Models; Nonprofessional Personnel; *Paraprofessional School Personnel; Performance; Performance Factors; Supervision; *Supervisors; *Supervisory Activities; *Supervisory Methods; Tutoring

ABSTRACT

Although increasing use of paraprofessionals to implement key program concepts has the advantage of increased availability and lower salaries, problems in maintaining acceptable levels of performance have also been reported. This study assessed the role of performance-related feedback on the work behavior of paraprofessional tutors in a remedial reading program. One randomly chosen tutor received publicly posted feedback each day on (1) degree of completeness with which one student's answers to comprehension-check questions were tutored; (2) accuracy with which that student's data sheet had been computed; and (3) time at which the first student's tutorial session had begun. The simple feedback package produced significant improvement in the measured level of the first two work-related behaviors, but not in the third. It was concluded that simple informational feedback and the objective measures of work performance on which it is based may be of considerable practical importance to supervisors in insuring satisfactory work performance by paraprofessionals in applied programs. (Author)

FILMED FROM BEST AVAILABLE COPY

ED 073387

Supervising Paraprofessionals: Performance-Related Feedback

James D. Barnard, Edward R. Christopherson
and Montrose M. Wolf

University of Kansas

CG 007 819

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Abstract

Applied researchers increasingly report the use of paraprofessionals to implement key program components, however, despite such apparent advantages as increased availability and lower salaries, problems in maintaining acceptable levels of on-the-job performance in these workers have also been reported. This study assessed the role of performance-related feedback on the work behavior of paraprofessional tutors in a remedial reading program. One randomly chosen tutor received publicly-posted feedback each day on 1) the degree of completeness with which one student's answers to comprehension-check questions were tutored; 2) the accuracy with which that student's data sheet had been computed; and 3) the time at which the first students' tutorial session had begun. The simple feedback package produced significant improvement in the measured level of the first two work-related behaviors, but not in the third, and subsequent increase in the number of tutor performances to contact feedback produced no additional improvement in this behavior. It was concluded that simple informational feedback (and the objective measures of work performance on which it is based) may be of considerable practical importance to supervisors in ensuring satisfactory work performance by paraprofessionals in applied programs.

Introduction

With the advent of a technology of applied behavior analysis, significantly large numbers of socially important problems have been investigated in the natural environment (Ulrich, Stachnik and Mabry, 1966). A characteristic trait, reported with increasing frequency by applied investigators, is the use of paraprofessional personnel to implement key components of applied programs (Staats, Minke & Butts, 1970; Allison, 1971). Despite the advantages of increased availability and lesser salaries, however, problems of maintaining adequate on-the-job performance with paraprofessionals have been reported (Pierce, 1971).

Performance-maintenance problems are relevant, particularly for programs employing paraprofessionals to assess the effects of intervention techniques that are carefully specified. Without some form of assurance that the procedures have been carried out, the investigator who obtains less than the expected experimental results, may erroneously attribute such outcomes only to the nature of the intervention package employed.

Given such problems, what controlling variables do applied supervisors have at their disposal to either prevent or ameliorate performance decrements in workers? One potential source of performance control is the variable known as feedback, or knowledge of results on performance. Simple informational feedback has been used by investigators to enhance performance in simulated work settings (Gibbs & Brown, 1955); actual industrial

settings (Feeney, 1972); institutional settings (Panyon, Boozer & Morris, 1970); and educational settings (Thomas, 1972, Thompson, Holmberg & Baer, 1971).

The present study investigated the effects of providing performance-related feedback on three work behaviors in a staff of paraprofessional reading tutors.

Footnote

This study is based upon a thesis submitted by the senior author to the Department of Human Development and Family Life, in partial fulfillment of the requirements of the Master of Arts Degree. The author wishes to recognize the co-authors for their role as mentors, and Ms's Carolyn Voss, Susan Rainey, Jane Guntert, and Mary Ann Mitchell for their contributions to the data collection process. This research was supported by a grant from the National Institute for Child Health and Human Development (HD 03144-03) to the Bureau of Child Research, University of Kansas, and a grant from the National Coordinating Center, National Program on Early Childhood Education (OEO-3-7-070706-3116) to the Kansas Center for Research in Early Childhood Education. Reprints may be obtained from the authors, Department of Human Development, University of Kansas, Lawrence, Kansas, 66044

METHOD

Subjects and Setting

Five paraprofessional reading tutors served as subjects. Of two adult women (ages 48 and 24), one had completed the sixth grade and the other had earned her GED high school equivalency diploma. These subjects worked a maximum of 20 hours per week at an hourly salary of $2.00. Three additional subjects were high school students employed through the Neighborhood Youth Corps. They were employed 10 hours per week at a salary of $1.60 per hour.

A remedial reading program served as the setting. The facility housed five small rooms for tutoring, each equipped with a desk, chairs, point-counter, cassette tape recorder, stop-watch and intercom. Students were tutored on a one-to-one basis for 30 minutes each week-day after school. While students read stories aloud, and answered questions over the materials, tutors monitored their performance, differentially awarding and subtracting points (Christophersen, Davis and Wolf, 1971), which could be exchanged for money at the session's end.

Three distinct classes of tutor behavior were examined. They were 1) completeness of tutoring, 2) accuracy of data sheet tabulations, and 3) the time tutoring was begun each day.

I Complete Tutoring

For complete tutoring, tutors had been trained to emit a short verbal praise statement following each correct answer.

A complete tutor response after a students' correct answer was thus defined as any audible utterance emitted by the tutor following that correct answer.

Following an incorrect answer, a complete tutor response was defined in terms of three components. A tutor-emitted verbalization was to occur, which directed the student to refer back to the reading selection, and which prompted the student to physically point-out, or read a question-related bit of information. For example, refer-back prompts usually took the form of "turn back to the story", and the like. Point out prompts, in contrast, were defined more specifically, and called for an observable physical response such as "put your finger on sentence three". A point-out was also said to have occurred whenever a student re-read a sentence aloud.

Recording of complete tutor behaviors was accomplished using a standardized data sheet, scoring each of the response components question by question, in an occurrence vs. non-occurrence fashion.

These data were combined, and expressed as a single percent complete tutor measure. This measure was obtained by dividing the number of completely tutored answers, both correct and incorrect, by the total number of answers given per session.

Interobserver reliability of this response measure was periodically assessed by having trained checkers independently analyze a single audio tape of a tutorial session. In addition, infrequent naive checker assessments were made by giving select visitors sets of written instructions along with the

requisite materials for making a check. Naive observer records were later compared to those produced by regular checkers.

Reliability of occurrence was assessed throughout, and was obtained by adding agreements between checkers on response occurrence, question by question within each response component column. An overall percent measure of occurrence agreement was then obtained by adding the total number of agreements and dividing by the number of agreements plus disagreements.

II Accuracy of Calculations

The accuracy with which tutors tabulated the daily written records of students oral reading and question answering performance was selected as the second dependent variable performance measure.

Two calculation products were obtained: the student's percent answer accuracy and his percent sentence reading accuracy. In order to obtain the student's percent answer accuracy the tutor was to count the number of correct answers; divide by the total number of questions answered; round off, and record the resultant on the data sheet. An identical sequence of operations was required in computing the oral reading accuracy measure.

At the close of each day, all data sheets were re-tabulated by a supervisor and ranked as being with or without errors. To be scored as errorless, each operation had to be correct. A single error, or dissagreement between tutor product and checker

product, was sufficient for the data sheet to be ranked as containing errors. Conversely, an errorless sheet was absolutely correct along each dimension. A standardized recheck sheet was used.

The percent of errorless data sheets served as the second daily performance measure, and was obtained by dividing the number of sheets classed as being errorless, by the total number of data sheets submitted on a given day. Reliability of this measure was assessed by having a second checker independently analyze and rank a day's data sheets. Agreement between checkers was defined as having occurred when two checkers ranked any one data sheet identically (as being with or without errors). The reliability index was obtained by dividing the number of agreements on categorization by the number of agreements plus disagreements. Naive checkers also provided additional checks on reliability of measurement via a program of written instruction similar to that described for measure one.

III Time Begin Tutor

The final variable chosen for analysis was the time at which each tutor actually began his first tutorial session on a given afternoon. Sessions were to have begun promptly at 3:30 p.m., and the beginning of a reading session was readily identifiable, as students always recited descriptive information aloud (their names, date, session number, etc.), prior to reading their first story selection. A session's beginning was thus defined as the time at which the student began this

recitation. Because each tutoring room was connected to a master-intercom system, observers had unobtrusive access to this response when it occurred. Regular observers listened for, and recorded start times on a standardized data sheet. Times were coded such that 3:30 p.m. was assigned a whole number 30; 3:42 a 42; 4:06 a 66, etc.. The mean time tutoring began was thus obtainable by summing the coded start times, and dividing by the number of observations.

Reliability was assessed by having a second observer listen for the beginning of a day's sessions, and record the times of occurrence. Observations within one minute of each other were required between the prime and reliability observers before agreement was credited with occurrence. The index of reliability was computed by summing the number of agreements on start-times and dividing by the number of agreements plus disagreements.

Design

A multiple baseline design (Baer, Wolf and Risley, 1968) was employed such that each of the three dependent variable measures was analyzed concurrently over time across all five tutors. The multi-component package of independent variables was then sequentially introduced to each of the dependent variables in a multiple time-series fashion (Campbell & Stanley 1968). Changes in levels of each treated variable were then assessed both relative to that variables' own baseline and relative to the concurrent baseline of the untreated variable.

Baseline data were collected over 21 working days. During this time, tutoring occurred unchanged as it had for roughly 10 months. Each day one randomly chosen tape per tutor was analysed in terms of complete tutor; all data sheets were checked for accuracy of calculations, and the session start times were recorded. Feedback on performance on each of these measures was withheld from tutors during baseline.

Feedback: Complete Tutor

A formal meeting of all reading program personnel was held to introduce feedback for complete tutoring. Written hand-outs were distributed, mentioning both the lifting of national wage price controls and the need for objective performance measures by which tutors could be evaluated for pay raises. The decision to select the complete tutor performance measure was announced.

Written definitions of, and instructions for completely tutoring were provided, as was a supplemental video tape which presented the most frequently encountered question types, and the appropriate models of complete tutor behavior for each.

Daily performance was made public each day by means of a feedback board, centrally displayed in the reading lab. This "score board" showed the randomly chosen tutor's name, his percent performance scores for each of the complete tutor components, and in addition contained a spot on which the project director's daily signature appeared. Tutor performance scores were also placed in the tutor's personal employee folder.

Feedback: Accuracy of Calculations

Baseline data were collected on the calculation measure for 29 days. On day 30, feedback was introduced to this variable. Tutors were again assembled as a group with written handouts distributed. The handout informed tutors that they would receive feedback on this performance through the feedback board. Component operations of the calculation process were also reviewed.

In this second condition, then, the randomly selected tutor not only received feedback on his complete tutor performance, but also received feedback on whether his data sheet was ranked as being with or without errors. This information was also introduced into the tutor's personal employee folder.

Feedback: Time Begin Tutor

Feedback was introduced to the time tutor variable on day 38 by calling a third brief group meeting. The time begin tutor variable was introduced as the final tutor performance measure to be monitored and posted on the feedback board. During this condition the randomly chosen tutor was given knowledge of his complete tutor percent scores, whether or not his calculations were done properly, and the actual time he began tutoring his first student on the day in question.

Supplementary Feedback

On day 133 a system was introduced which provided tutors with feedback on variables 2 and 3, supplemental to that already

provided via the display board. In this condition, each tutor received a daily feedback sheet which listed all data sheets computed, and whether or not each was marked as being with or without errors. The sheet also listed the tutor's starting time and whether he was judged as being punctual or tardy. These daily sheets were signed by tutors, returned to the supervisor and placed in the personal employee folders.

RESULTS

Reliability

Regular observers checked the reliability of the complete tutor measure on 28 sessions resulting in a mean percent agreement of 98. Naive observer agreement averaged 84 percent. The reliability of the accuracy of calculation measure was likewise assessed on 20 occasions with agreement between checkers averaging 97 percent. Mean agreement between naive and regular checkers was 89 percent. Independent observer records for 25 sessions were compared for the time begin tutor variable with percent agreement between checkers averaging 9%.

Insert Figure 1 About Here

Figure 1 summarizes the mean performance of all five tutors on each of the three dependent measures over 148 regular working days. Due to the relative stability of the observed condition trends, sessions 61 through 122 inclusive, have been uniformly deleted from the graphic summarization of the data. Analysis of the condition means shows that the mean percent with which students' answers were completely

tutored increased from a baseline mean of 42% to a feedback mean of 93%. The mean percent of errorless data sheets shifted upward, although in a less pronounced fashion, from a baseline mean of 62 to a feedback mean of 75. The mean time at which tutoring was begun averaged 3:43 pm during baseline and 3:49 pm in the subsequent feedback condition.

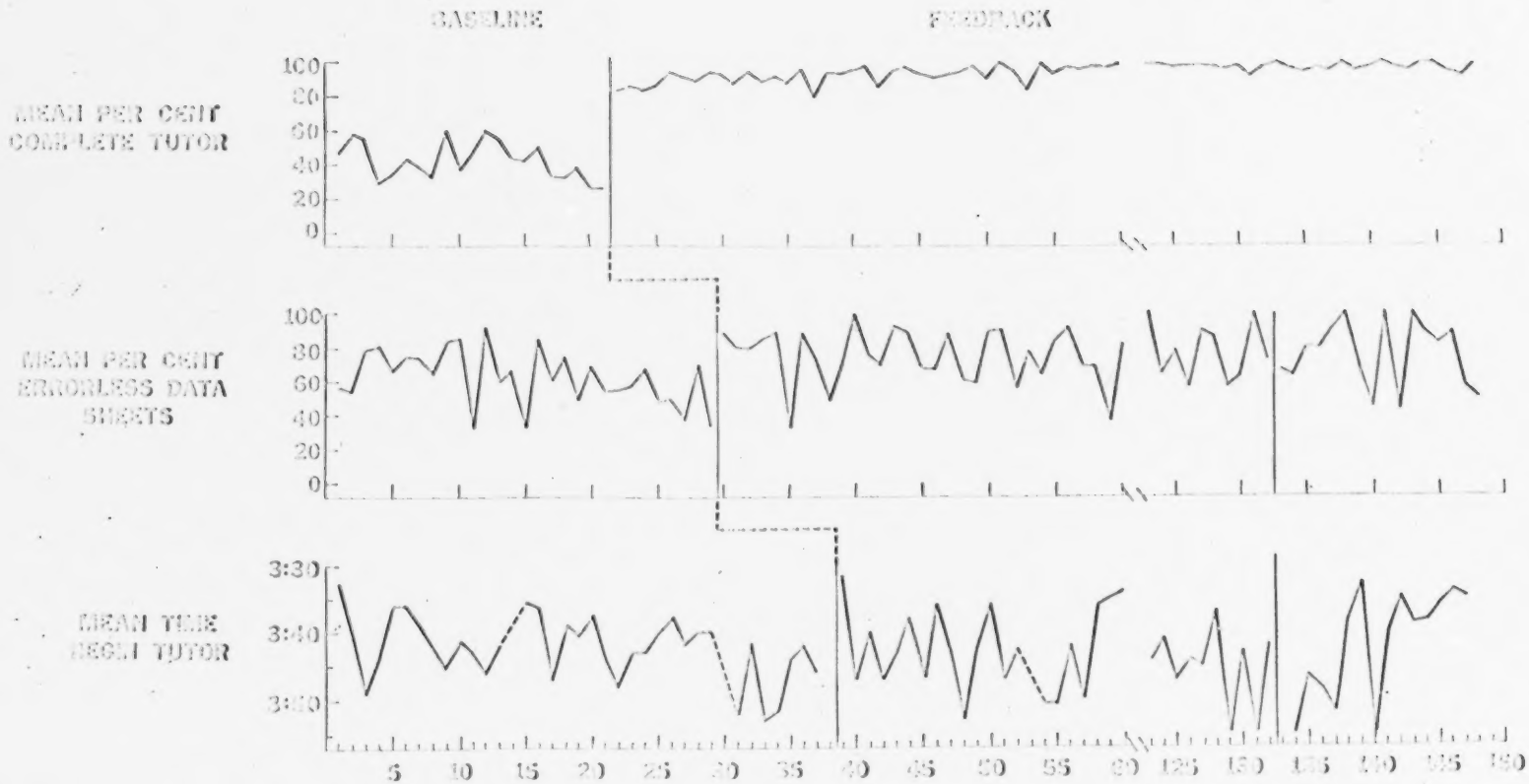
Significance Analyses

Five t-tests were performed to asses the probability with which the observed differences in condition means could be attributed to chance variation. Data were treated as correlated and observations were paired (Dixon and Massey, 1957). Three of these t-tests compared the final eight days of baseline with the first eight days of feedback for each variable. Resultant t-scores showed the condition mean difference to be significant at the .001 level for variable one; significant at the .01 level for variable two; and significant at the .05 level for variable three. Two additional t-tests analysed the natural variation in the final 16 days of baseline for variables two and three, and thus served as controls for simultaneous changes which occurred in variables one and two with the introduction of feedback. Both of these t-scores were found to be significant.

Supplementary Feedback

Inspection of the data from sessions 123-148 shows the insignificant effects of introducing additional feedback to variables two and three. The mean percent of errorless data sheets shifted from 77 to 75, and a slight reduction in the mean time begin tutor variable from 3:45 to 3:42 pm.

Figure 1 A record of mean tutor performance (ordinate) over session days (abscissa), showing mean percent complete tutor, mean percent errorless data sheets, and mean time begin tutor. Heavy upright lines indicate offset of baseline and onset of feedback conditions. Broken lines connecting data points represents estimates of performance levels on days when actual data points were unavailable. Parallel hash marks on abscissa indicate that data are available, but omitted from the graphic summarization.


BASELINE
FEEDBACK
MEAN PER CENT COMPLETE TUTOR
100
80
60
40
20
0
MEAN PER CENT ERRORLESS DATA SHEETS
100
80
60
40
20
0
MEAN TIME REQD TUTOR
3:30
3:40
3:50
5
10
15
20
25
30
35
40
45
50
55
60
125
130
135
140
145
150

DISCUSSION

The unsatisfactory baseline levels of crucial on-the-job performances documented here and elsewhere suggest a need for the definition of objective work performance measures along with the periodic assessment of the performance levels obviated by such measures.

These multiple baseline data suggest a functional relationship between the introduction of the complex of treatment variables and the observed behavioral changes. While it is tempting to attribute these changes to the feedback component alone, such a conclusion would be premature, as no attempt was made in this analysis to isolate the feedback component from the additional confounding components (announcement of the renewed possibility of pay raises, the written hand-outs, etc.). The most tenable conclusion would be that this combination of variables effected significant positive changes in two of the variables to which it was introduced. In addition, it was concluded that the additional time required to provide tutors with supplementary feedback yielded a negligible return, and that the mere presence of feedback seemed equally as crucial as the amount or degree of complexity of feedback.

To the extent that the changes in performance could be attributed primarily to the effects of feedback, these data would add to the generality of the findings of others who have reported feedback to be a powerful source of control on performance (Gibbs & Brown, 1955; Thomas, 1972; Thompson, Holmberg & Baer 1971; Feeney 1972; Panyon Doozer & Morris, 1972). In

addition, a relationship seems evident between simple informational feedback, and the nature of the task variable to which it is presented. An interaction effect seems plausible, and may account for both the magnitude and longivity of the effect to be realized with feedback. Such a differential effectiveness function has been suggested by Annette (1969).

Most important to consider is that the observed behavioral changes were effected very practically. The mean amount of time required to sample any one tutor, then rate and post him or her performance was only 28 minutes per day. The gains, however, were significant to the extent that the original reading research undertaking could be continued with the supervisor's assurance that the specified procedures were being reliably implemented.

Finally, these results should suggest that while incapable of being viewed as a supervisory panacea, simple feedback on work performance might well be more widely considered by supervisors of paraprofessionals as a practical, effective source of performance control. Equally important, the reliable use of non-degreed and/or less formally educated personnel might be greatly facilitated by arranging supervisory conditions such that they are capable of both generating and feeding back objective information about work performance to these paraprofessionals. Systematic replication of these and other findings however, would seem necessary to establish the generality of the control exerted by various types of feedback accross a multiplicity of work performances, settings, and workers.

BIBLIOGRAPHY

Annette, J. Feedback and human behavior. Baltimore: Penguin Books, 1969.

Baer, D.M., Wolf, M.M., and Risley, T.R. Some current dimensions of applied behavior analysis. Journal of Applied Behavior Analysis, 1968, 1, 91-97.

Campbell, D.T., and Stanley, J.C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.

Christopherson, E.R., Davis, K.E., and Wolf, M.M. A tutoring program employing paraprofessionals. Lawrence, Kansas (Kansas RRI), Winter, 1972.

Dixon, W.J., and Massey, F.J. Introduction to statistical analysis (2nd ed.) New York: McGraw-Hill, 1957.

Allison, D.G. Model programs: Programmed tutorial reading project, Indianapolis, Indiana, U.S. Government Printing Office, Washington, D.C., 1971, Catalogue No. HE 5.230:30030, Stock No. 1780-1781, National Center for Educational Communication, U.S. Department of Health Education and Welfare, Office of Education.

Feeney, E.J. New tool: Reinforcement for good work. Business week magazine, December 16, New York: McGraw-Hill, 1972.

Gibbs, C.B. and Brown, R.D. Increased production from the information incentive in a repetitive task. Nature, 1972, 24, 374-379.

Kenyon, K., Bolzen, T. and Morris, K. Feedback to attendants as a reinforcer for applying operant techniques. Journal of Applied Behavior Analysis, 1970, 3, 1-4.

Kiesee, C.K. Experimental recreation in an urban poverty area. Unpublished doctoral dissertation, University of Kansas, 1971.

Staats, A.W., Minke, K.A. and Butts, P. A token-reinforcement remedial reading program administered by instructional technicians. Technical Report No. 127, May, 1970, Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, Center No. C-03, Contract No. OE-5-10-154, Department of Health Education and Welfare, Office of Education.

Thomas, D.R. Self-monitoring as a technique for modifying teaching behavior. (Doctoral dissertation, University of Illinois at Urbana-Champaign) Ann Arbor, Michigan: University Microfilms, 1972.

Thompson, C.L., Wolmberg, R.G., and Baer, D.M. The experimental analysis of training procedures for pre-school vocations. Paper presented as a symposium at the Biennial meeting of the Society for Research in Child Development, Minneapolis, 1971.

Ulrich, R., Stachnik, T. and Mabry, J. (eds.) Control of human behavior Glenview, Illinois: Scott, Foreman, 1966.